Feedback Control and Optimization in Online Advertising

Niklas Karlsson, Ph.D.
Vice President Algorithm Research
Advertising.com Group, R&D

December 6, 2011
Agenda

- Motivation
- Evolution of Online Ad Optimization Systems
- Feedback Control of Ad Campaigns
- Challenges
- Exploration & Exploitation
- Experimental Results
- Conclusions
Online Display Advertising – The Opportunity

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TV</td>
<td>$180,460</td>
<td>$185,788</td>
<td>$172,320</td>
<td>$174,836</td>
<td>$183,177</td>
</tr>
<tr>
<td>Newspapers</td>
<td>$130,178</td>
<td>$123,109</td>
<td>$102,136</td>
<td>$97,703</td>
<td>$97,228</td>
</tr>
<tr>
<td>Internet</td>
<td>$40,242</td>
<td>$49,544</td>
<td>$54,087</td>
<td>$60,253</td>
<td>$68,557</td>
</tr>
<tr>
<td>Magazines</td>
<td>$59,196</td>
<td>$65,588</td>
<td>$45,415</td>
<td>$42,762</td>
<td>$42,573</td>
</tr>
<tr>
<td>Radio</td>
<td>$38,583</td>
<td>$37,630</td>
<td>$33,647</td>
<td>$33,280</td>
<td>$34,216</td>
</tr>
<tr>
<td>Outdoor</td>
<td>$31,752</td>
<td>$31,888</td>
<td>$29,112</td>
<td>$29,828</td>
<td>$31,430</td>
</tr>
<tr>
<td>Cinema</td>
<td>$2,268</td>
<td>$2,377</td>
<td>$2,180</td>
<td>$2,274</td>
<td>$2,422</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>$482,680</strong></td>
<td><strong>$486,924</strong></td>
<td><strong>$438,896</strong></td>
<td><strong>$440,936</strong></td>
<td><strong>$459,603</strong></td>
</tr>
</tbody>
</table>

*Note: currency conversion at 2008 average rates*
*Source: ZenithOptimedia as cited in press release, October 19, 2009*

Online spending is the only one increasing
Online Display Advertising – The Opportunity

Online spending is the only one increasing (and rapidly)
Online Display Advertising – The Opportunity

Share of media time

Share of advertising budget

There is still a lot of room for growth
Stats of the Ad.com Ad Optimization System

Currently running 24/7/365 in 4 countries: US, Canada, UK, and Japan

Up to 3 Billion Impressions per day
  • Over 50,000 impressions per second at peak time

In 2010:
  • 500 billion impressions
  • 60 million conversions

Lifetime totals:
  • 5 trillion impressions
  • 400 million conversions
The 10,000 Foot View

Publisher Constraints:
- Payment model – may charge per impression, click, or conversion
- Allowability – may prohibit certain types of ads to be displayed

Advertiser Constraints:
- Payment model – may pay per impression, click, or conversion
- Allowability – may restrict on what web sites to be served
- Targeting – may only want to be shown to internet users in a certain geo location, or from a specific demographic
- Frequency – may limit how often the same user is shown the ad
- Campaign Delivery:
  - The total ad budget may have to be delivered according to a plan
  - The served impressions may have to generate no less than a prescribed click-through or conversion rate

When, where, and to whom should we show an ad?
The Evolution of Online Ad Optimization Systems

The State in 1998
- Few advertisers and publishers
- Few companies offer optimization solutions
- Advertisers are unsophisticated
- User data is virtually non-existent

The State in 2011
- Large number of advertisers and publishers
- Many companies offer optimization solutions
- Advertisers are savvy
- User data is highly available
The Evolution of Online Ad Optimization Systems

Generation 1 – “Centralized and Open Loop”:

\[
\max_{\alpha(t, i, j)} \int_0^\infty \sum_i \sum_j \alpha(t, i, j) n_1(t, j) p(t, i, j) \eta(i) \, dt
\]

subject to (e.g.)

\[
n_\Lambda(t, i) \leq c_i(t), \quad i \in \text{\forall ads}
\]

Algorithm features:

- Forecasts \( n_1(t, i, j) \) and \( p(t, i, j) \)
- Maximizes network revenue by centrally solving the constrained optimization problem

Notation:

- \( n_1(t, j) \) rate of impressions to site \( j \)
- \( n_\Lambda(t, i, j) \) rate of conversions for ads \( i \) displayed on site \( j \)
- \( p(t, i, j) \) impression-to-conversion probability for ad \( i \) shown on site \( j \)
- \( \eta(i) \) network revenue when ad \( i \) converts
- \( \alpha(t, i, j) \) relative allocation of impressions on site \( j \) to ad \( i \)
The Evolution of Display Advertising Systems

Generation 2 – "De-centralized and Feedback-based":

\[
\max_{u(.)} \int \sum_{i} \sum_{j} f \left( u(t,i), p(t,i,j) \eta(i) \right) dt
\]

subject to (e.g.)

\[ n_A(t,i) \leq c_i(t), \quad i \in \forall \text{ads} \]

Algorithm features:

- Forecasts \( p(t,i,j) \) (the revenue per conversion \( \eta(i) \) is given)
- Satisfies campaign delivery constraints via decentralized error feedback control
- Implements ad serving as an impressions-based auction exchange

Notation:

- \( u(t,i) \) control signal/bid adjustment for ads bidding on site j impressions
- \( f(u, p\eta) \) bid adjustment function
The Evolution of Display Advertising Systems

Generation 3 – “User-segmentation”:

\[
\max_u \int_t \sum_{ad} \sum_{site} \sum_{userseg} \sum_k f \left( u(t, i), p(t, i, j, k) \eta(i) \right) dt
\]

subject to (e.g.)

\[ n_A(t, i) \leq c_i(t), \quad i \in \forall ads \]

Algorithm features:

- Forecasts \( p(t, i, j) \) (the revenue per conversion \( \eta(i) \) is given)
- Satisfies campaign delivery constraints via decentralized error feedback control
- Implements ad serving as an impressions-based auction exchange
- Incorporates user segments as a dimension for optimization

Notation:

- A user segment may represent e.g. a user’s geographic location, age, income bracket, hobbies, and/or frequency of visiting a specific site
The Evolution of Display Advertising Systems

Generation 4 – “Adaptive user-segmentation”:

$$\max_{u(.)} \int \sum_{t} \sum_{ad} \sum_{i} \sum_{site} \sum_{j} \sum_{userseg} \sum_{k(t)} f \left( u(t, i), p(t, i, j, k(t)) \eta(i) \right) dt$$

subject to (e.g.)

$$n_A(t, i) \leq c_i(t), \quad i \in \mathcal{V}_{ads}$$

Algorithm features:

- Forecasts $p(t, i, j)$ (the revenue per conversion $\eta(i)$ is given)
- Satisfies campaign delivery constraints via decentralized error feedback control
- Implements ad serving as an impressions-based auction exchange
- Incorporates user segments as a dimension for optimization
- Adaptively updates segment definitions

Notation:

$k(t)$ a user segment that is re-defined over time
The Evolution of Display Advertising Systems

**Generation 5 – “Advertiser Cost Minimization”:**

\[
\min_{u(\cdot, i^*, \cdot)} \max_{u(\cdot, \cdot, \cdot)} \int_t \sum_{j} \sum_{\text{userSeg}} g_{i^*} (\ldots, u(t, i, j, k), p(t, i, j, k(t)), \ldots) \, dt
\]

subject to (e.g.)

\[n_A(t, i) \leq c_i(t), \quad i \in \forall \text{ads}\]

**Algorithm features:**

- Forecasts \(p(t, i, j)\)
- Satisfies campaign delivery constraints via decentralized error feedback control
- Implements ad serving as an impressions-based auction exchange
- Incorporates user segments as a dimension for optimization
- Adaptively updates segment definitions
- Defines the optimization problem as a non-cooperative multi-player game

**Notation:**

\(g_{i^*}(\cdot)\) cost model dictating how much an advertiser \(i^*\) pays for an impression \(i,j,k\) in a competitive market place
The Evolution of Display Advertising Systems

Traffic & Response Rate Predictors

Centralized Optimization

Ad Allocation Schedule

Actual Advertisement Result

Constraints

Ad Serving/Internet

Campaign Delivery Constraints

Campagne Management

Ad placement valuation

Max Bids

Bids

Advertisement Result

Campaign Controller

Ad Serving/Internet (market place)

Generation 5
• De-centralized (scales well)
• Fine segmentation (superior yield)
• Closed loop (learns from mistakes)
• Campaign-centric (objectives are aligned)

Generation 1
• Centralized (does not scale well)
• Coarse segmentation (inferior yield)
• Open loop campaign delivery (does not learn from mistakes)
• Network-centric (network & campaign objectives mis-aligned)
Feedback Control of Ad Campaigns

Deliver the campaign as fast as possible while satisfying

\[
\begin{align*}
    n_A(t, i) &\leq c_{1,i}(t) \quad \text{(Smoothness)} \\
    n_A(t, i)/n_1(t, i) &\geq c_{2,i}(t) \quad \text{(Performance, Gen. 1-4)} \\
    n_A(t, i, j) &\leq c_{3,i}(t) \quad \text{(Spread)} \\
    n_A(t, i, k) &\leq c_{4,i}(t) \quad \text{(Partition)} \\
    \min p(t, i, j, k)b(t, i, j, k^*) &\leq c_{5,i}(t) \quad \text{(Cost, Gen. 5)}
\end{align*}
\]
Challenges for Optimization

Important aspects to consider during the design of a campaign management system are

**Very small event rates:** On the order of one click/action per thousands or millions of impressions; very noisy observed rates

**Sparse data:** We have little or no historical data for most ad, site, user-segment combinations

**Latency:** Actions may occur long after the ad was shown (sometimes days or even weeks later)

**Time variability:** Volume and response rate are subject to trends, seasonality, and sudden step changes

**Coupling effects:** A campaign entering or exiting the network impacts the delivery of other campaigns
Challenges – Latency

Definition
- Latency is the time between when an impression is served to a user and the user converts (e.g. make a purchase) – assuming a conversion occurs

Observation
- The latency distribution varies dramatically from one ad campaign to another
Challenges – Time-varying Volume and Response Rate

Definition

- Impression Volume is the total number of impressions awarded to a campaign in a given time window
- Click-through-rate is the ratio of impressions for which the user clicked on the advertisement

Observation

- The impression volume and click-through-rate exhibit a noisy time-varying behavior with a time of day component, a slowly varying component, and occasional jumps
Challenges – Coupling Effects

Facts
- The highest bidding campaign on any given segment is awarded any desired and available inventory on this segment.
- The price-volume curve is non-smooth and varies across campaigns.
- Small changes in the control signal may result in jumps in the revenue rate.

Definition
- Coupling effects is the interdependency among campaigns in the network.

Observation
- Coupling effects in an auction exchange lead to a discontinuous plant.
Exploration versus Exploitation

Online ad optimization with advertisers entering and exiting the network is naturally viewed as a *Multi-armed restless bandit problem*.

A sound valuation strategy

- Minimizes the long term cost for a campaign by carefully balancing information acquisition (exploration) and capitalization of all learning (exploitation)
- Account for the risk and opportunity associated with the uncertainty in estimates

Maximum bid = Expected Revenue Per Impression + Learning Value
Real campaign – example 1

This example illustrates
- Behavior with old and new campaign delivery solution
- Budget delivery control with acceleration/deceleration
This example illustrates

- Step response in budget delivery when the desired pacing is adjusted
- (Re)-Initialization Control

**Campaign Information**

**General**
- Name: Sample Campaign 2
- Owner:
- Instance ID:
- Campaign Start: 15-OCT-2008
- Controller Start: 31-MAR-2009
- Payment Type: USD 1.85 CPM
- Performance Type: NONE
- Total Budget: USD 260,990
- Reference Budget: USD 0

**Campaign Status**
- Bidding Status: Active

**Controller**
- Name:
- Method: Conventional Smoothness
- Revenue Smoothness: Price
- Performance: Price

**Delivery Objective:** Defined Schedule

**Smoothness delivery with step change in desired delivery**

**Definit Schedule Tracking**

**Expressions (Observed Hourperiod)**

<table>
<thead>
<tr>
<th>DATE</th>
<th>TYPE</th>
<th>OLD</th>
<th>NEW</th>
</tr>
</thead>
<tbody>
<tr>
<td>26-FEB-2009 18:45</td>
<td>PCAPM</td>
<td>3.5</td>
<td>1.1</td>
</tr>
<tr>
<td>27-FEB-2009 03:15</td>
<td>Initialization</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>29-FEB-2009 20:30</td>
<td>PCAPM</td>
<td>2.8</td>
<td>3.5</td>
</tr>
<tr>
<td>10-FEB-2009 16:45</td>
<td>Initialization</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18-FEB-2009 16:06</td>
<td>Budget</td>
<td>23097.28</td>
<td>260995.67</td>
</tr>
</tbody>
</table>
Click-through-rate is improving

Performance control is increasing

This example illustrates

- Simultaneous Budget delivery & Performance Control
- Interaction between Smoothness and Performance Control
Beyond the Closed Network

The presence of a market price of any given impression enable

- External bidders to compete for impressions in our network
- Internal ad campaigns to bid for impressions in other exchange-based networks
Conclusions

We established that

- Internet advertising is a big and rapidly growing industry
- Optimal online advertising is an extremely high-dimensional problem with additional challenges due to tiny event rates, latency, time variability, and coupling effects
- Decentralized feedback control has proven to be of great value for controlling ad campaigns

We did (among other things) not cover

- Business requirement gathering process
- Research, development, testing, and deployment process